Hot Workshops

DOE Office of Science Data Management Workshops

Richard P. Mount

SLAC

National Collaboratories Meeting

August 10, 2004

DM Workshops

1. SLAC, March 16-18, 2004
   - ~100 Participants
   - Focus on Needs, Technology, Gaps
2. SLAC, April 20-22, 2004
   - ~30 participants
   - Focus on common understanding of how to structure a report
3. Chicago, May 24-26, 2004
   - ~80 participants
   - Focus on gathering input to report.

Overview (1)

- Science, like business, national security and even everyday life, is becoming more and more data intensive.
- In many sciences the data-management challenge already exceeds the compute-power challenge in its needed resources.
- Leadership in applying computing to science will necessarily require both world-class computing and world-class data management.

Overview (2)

- Data and Information threaten to overwhelm:
  – Simulation-driven science;
  – Experiment/Observation-driven science;
  – Information-intensive science.
- DOE ASCR has been doing relatively little about this.
- Some “application sciences” are doing much more, but in an largely uncoordinated just-enough-to-survive mode.

Office of Science Data Management Workshops One View

- Determine application science needs;
- Determine what is, and will be available from technology and computer science;
- Document the Gaps;
- Design a Program that will be able to address current and future gaps.

Office of Science Data Management Workshops Another View

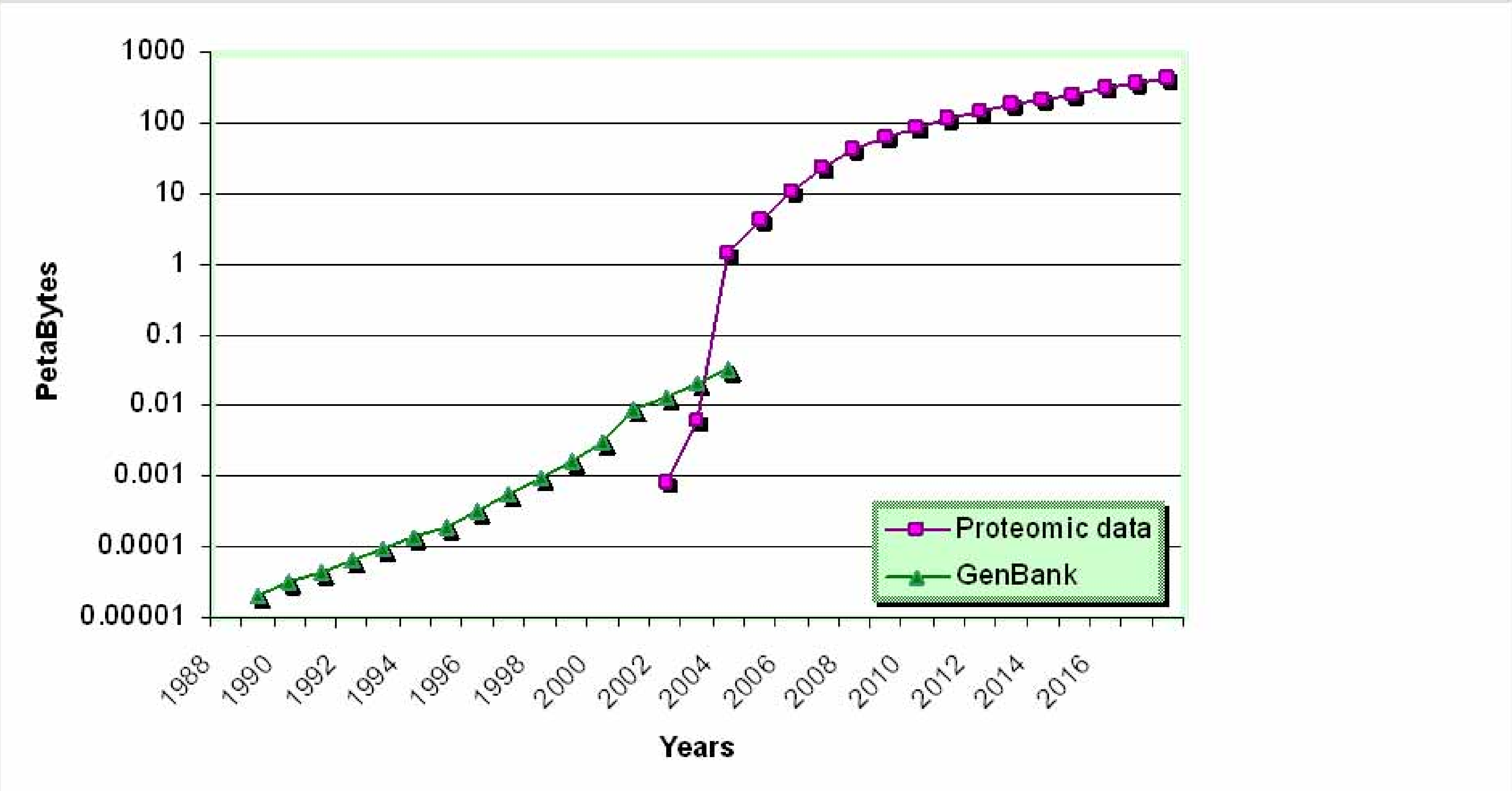
- Data-challenged scientists find out they are not alone – at least they have each other;
- Data-challenged scientists have problems understanding CS gobbledygook
  (ontology – AARRGGHHH);
- Everyone realizes that its both fun and important to bridge this gap;
- CS types have lots of fun promoting all their ideas;
- Strong consensus for a new program with emphasis on application-CS collaboration.

Workshop 1
The Data-Challenged Perspective

High throughput biology data management and data intensive computing drivers

George Michaels

Growth of Proteomic Data vs. Sequence Data



Community Data Resource

R & D Challenges

Design and Integration of the major databases

Huge data volumes, great schema complexity - need for new types of databases (hardware and software)

Database technologies – object-relational, graph DBs, …

Data standards, representations, ontologies for very complex objects

User Access Systems for browsing, query, visualization, and to run analysis or simulations

Supporting Simulation from DBs - how to allow users to utilize models and run simulations; how to link simulations to underlying data

Integration
- Provide integrated view of the biology
- With data from other community sources.

Community access to compute power to run long time-scale simulations

IP issues and reward system

How to represent **incomplete, sparse, conflicting data**

MFE Simulation Data Management

SLAC DMW 2004

March 16, 2004

W. W. Lee and S. Klasky
Princeton Plasma Physics Laboratory
Princeton, NJ

Data Management challenges

- GTC is producing TBs of data
  - Data rates: 80Mbs now, 1.6Gbs 5 years.
  - Need QOS to stream data.
- This data needs to be post-processed
  - Essential to parallelize the post-processing routines to handle our larger datasets.
  - We need a cluster to post process this data.
    - M (supercomputer processors) x N (cluster processors) problem.
    - QOS becomes more important to sustain this post-processing.
- The post-processed data needs to be shared among collaborators
  - Different sections of the post-processed data may go to different users .
  - Post-processed data, along with other metadata should be archived into a relational database.

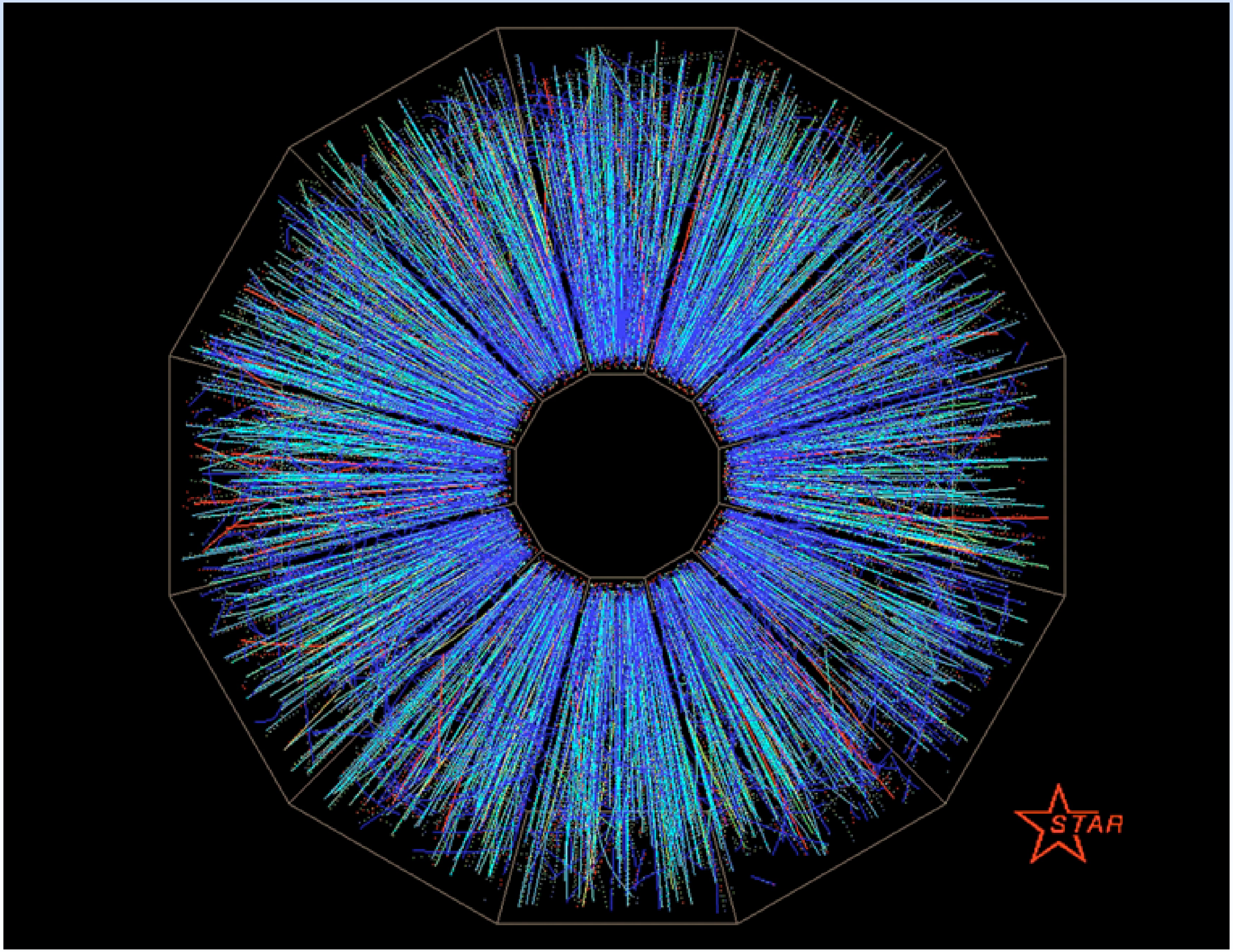
PPPL PRINCETON PLASMA PHYSICS LABORATORY

Nuclear Physics Data Management Needs

Bruce G. Gibbard

SLAC DMW2004 Workshop
16-18 March 2004

Digitized Event In STAR at RHIC



Data Volumes in Current RHIC Run

- Raw Data (PHENIX)
  - Peak rates to 120 MBytes/sec
  - First 2 months of '04, Jan & Feb
    - 10^9 Events
    - 160 TBytes
  - Project ~ 225 TBytes of Raw data for Current Run
- Derived Data (PHENIX)
  - Construction of Summary Data from Raw Data then production of distilled subsets from that Summary Data
  - Project ~270 TBytes of Derived data
- Total (all of RHIC) = 1.2 PBytes for Current Run
  - STAR = PHENIX
  - BRAHMS + PHOBOS = ~ 40% of PHENIX

NP Analysis Limitations (2)

- It seems that it is less frequently possible to do Topological Analyses in NP than in HEP so Statistical Analyses are more often required
  - Evidence for this is rather anecdotal – not all would agree
  - To the extent that it is true, final analysis data sets tend to be large
  - These are the data sets accessed very frequently by large numbers of users … thus exacerbating the data management problem
- In any case the extraction and the delivery of distilled data subsets to physicists for analysis currently most limits NP analyses

Combustion Science Data Management Needs

Jacqueline H. Chen
Combustion Research Facility
Sandia National Laboratories
jhchen@sandia.gov

DOE Data Management Workshop
SLAC
Stanford, CA
March 16-18, 2004

Sponsored by the Division of Chemical Sciences Geosciences, and Biosciences, the Office of Basic Energy Sciences, the U. S. Department of Energy

CRF

Data management challenges for combustion science

- 2D complex chemistry simulations today: 200 restart files $(x, y, Z_1, \ldots Z_{50})$ skeletal n-heptane 41 species, 2000x2000 grid, 1.6 Gbytes/time x200 files = 0.32 Tbyte, 5 runs in parametric study 1.6 Tbytes raw data
- Processed data: 2 Tbyte data

- 3D complex chemistry simulations in 5 years: 200 restart files $(x, y, Z_1, \ldots Z_{50})$ skeletal n-heptane 41 species, 2000x2000x2000 grid, 3.2 Tbytes/time x 200 files = 640 Tbytes per run, 5 runs = 3.2 Petabytes raw data
- Processed data: 3 Petabytes
- Combustion regions of interest are spatially sparse
- Feature-borne analysis and redundant subsetting of data for storage
- Provenance of subsetted data
- Temporal analysis must be done on-the-fly
- Remote access to transport subsets of data for local analysis and viz.

Why Feature Tracking?

- Reduce size of data
  - How do you find small ROI's in a large 3D domain?
  - Retrieve and analyze only what you need
- Provide quantification
  - Can exactly define ROI chosen & do specific statistics
- Enhance visualization
  - Can visualize features individually
  - Can color code features
- Facilitate event searching
  - Events are feature interactions

High Energy Physics Data Management

Richard P. Mount
Stanford Linear Accelerator Center

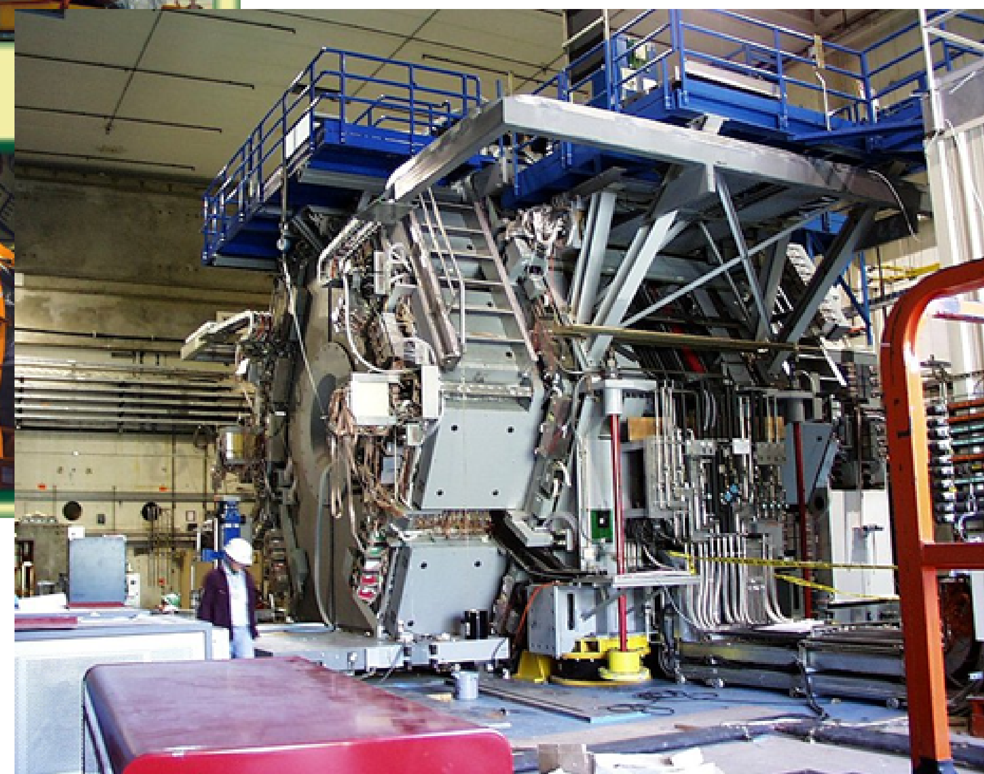
DOE Office of Science Data Management Workshop,
SLAC March 16-18, 2004

BaBar Experiment at SLAC

Taking data since 1999.

Now at 1 TB/day rising rapidly

Over 1 PB in total.



Matter-antimatter asymmetry

Understanding the origins of our universe

CMS Experiment: "Find the Higgs"
~10 PB/year by 2010

Home | Search | SiteMap | Contact Us| Credits

About CERN | Visit CERN | Education | Spotlight on ... | Ask an Expert

Spotlight on...

Home> Spotlight on ...

CMS - The Compact Muon Solenoid

The Compact Muon Solenoid is one of the experiments for the Large Hadron Collider at CERN. The collaboration involves about 1870 scientists coming from 150 Institutions distributed in 31 Nations.

Characteristics of HENP Experiments 1980 – present

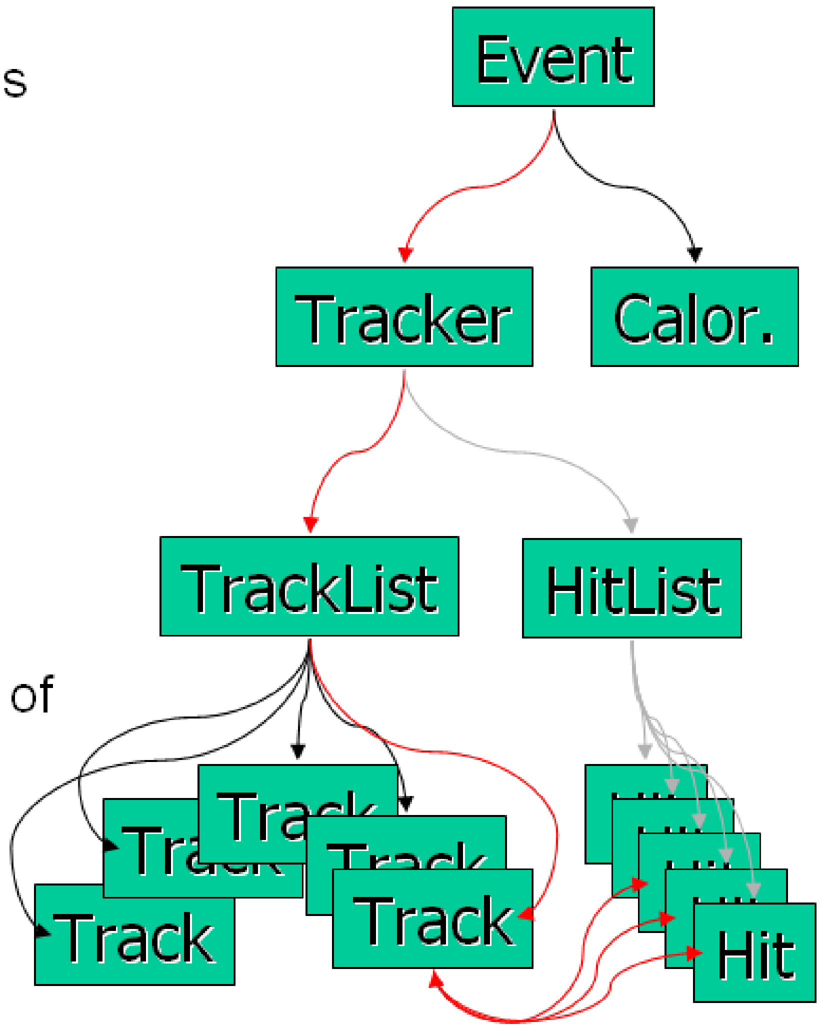
Large, complex detectors	$\Rightarrow$	Large, (approaching worldwide) collaborations: 500 – 2000 physicists
	$\Rightarrow$	Long (10 – 20 year) timescales
God *does* play dice	$\Rightarrow$	High statistics (large volumes of data) needed for precise physics

$\Rightarrow$ Typical data volumes: 10000*n* tapes

$(1 \leq n \leq 20)$

HEP Data Models

- HEP data models are complex!
  - Typically hundreds of structure types (classes)
  - Many relations between them
  - Different access patterns
- Most experiments now rely on OO technology
  - OO applications deal with networks of objects
  - Pointers (or references) are used to describe relations



Dirk Düllmann/CERN

Today's HENP Data Management Challenges

- Sparse access to objects in petabyte databases:
  - Natural object size 100 bytes – 10 kbytes
  - Disk (and tape) non-streaming performance dominated by latency
  - Approach - current:
    - Instantiate richer database subsets for each analysis application
  - Approaches – possible
    - Abandon tapes (use tapes only for backup, not for data-access)
    - Hash data over physical disks
    - Queue and reorder all disk access requests
    - Keep the hottest objects in (tens of terabytes of) memory
    - etc.

Data Requirements for Climate and Carbon Research

David C. Bader
Chief Scientist, DOE Climate Change Prediction Program
Presentation Material Courtesy of:
John Drake, ORNL
Don Middleton, National Center for Atmospheric Research
and
CCPP Scientists

UCRL-PRES-202932

ESG: Challenges

- Enabling the simulation and data management team
- Enabling the core research community in analyzing and visualizing results
- Enabling broad multidisciplinary communities to access simulation results

We need integrated scientific work environments that enable smooth WORKFLOW for knowledge development: computation, collaboration & collaboratories, data management, access, distribution, analysis, and visualization.

PCMDI

CCPP

Managing Data for the World Wide Telescope aka: The Virtual Observatory

Jim Gray

Alex Szalay

SLAC Data Management Workshop

Information Avalanche

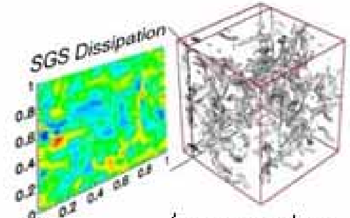
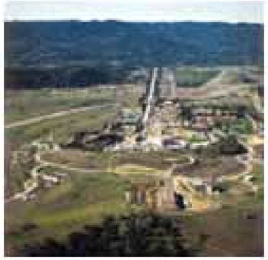


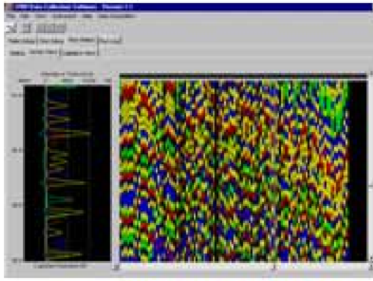
Image courtesy
C. Meneveau & A. Szalay @ JHU

- In science, industry, government,….
  - better observational instruments and
  - and, better simulations

  producing a data avalanche
- Examples
  - BaBar: Grows 1TB/day
    2/3 simulation Information
    1/3 observational Information
  - CERN: LHC will generate 1GB/s .~10 PB/y
  - VLBA (NRAO) generates 1GB/s today
  - Pixar: 100 TB/Movie



BaBar, Stanford



P&E Gene Sequencer From
http://www.genome.uci.edu/

- **New emphasis on informatics:**
  - **Capturing, Organizing, Summarizing, Analyzing, Visualizing**

Space Telescope

- Download (FTP and GREP) are not adequate
  - You can GREP 1 MB in a second
  - You can GREP 1 GB in a minute
  - You can GREP 1 TB in 2 days
  - You can GREP 1 PB in 3 years.
- Oh!, and 1PB ~3,000 disks
- At some point we need
  - **indices** to limit search
  - **parallel** data search and analysis
- This is where databases can help
- Next generation technique: **Data Exploration**
  - Bring the analysis to the data!

The Speed Problem

- Many users want to search the whole DB ad hoc queries, often combinatorial
- Want ~ 1 minute response
- Brute force (parallel search):
  - 1 disk = 50MBps => ~1M disks/PB ~ 300M$/PB
- Indices (limit search, do column store)
  - 1,000x less equipment: 1M$/PB
- Pre-compute answer
  - No one knows how do it for all questions.

Report Highlights

Existing Data-Management Funding

Project or Activity	DOE CS Funding $M/year	DOE Application-Science Funding $M/year	Data-Management as fraction of total
SciDAC: Scientific Data Management ISIC	3.0		100%
SciDAC: Particle Physics Data Grid	1.5	1.7	30%
SciDAC: High-Performance Data Grid Toolkit	0.8		50%
SciDAC: DOE Science Grid	1.9		10%
SciDAC: Fusion Collaboratory	1.8		20%
SciDAC: Earth System Grid II	1.8	0.4	100%
SciDAC: Logistical Networking1	0.4		70%
Collaboratory for Multi-Scale Chemical Science	1.9		65%
Storage Resource Management for Data Grid Applications	0.5		100%
Scientific Annotation Middleware	0.6		100%
Astronomy and Astrophysics		0.5	100%
Biology		2.0	100%
Climate		4.0	100%
Chemistry/Combustion		0.5	100%
Fusion		0.5	100%
High Energy Physics		5.0	100%
Nuclear Physics		1.0	100%
Nanoscience		0.5	100%
TOTAL Existing Activity	8.82	14.91	100%

Computer Science Areas

1 Workflow, dataflow, data transformation

1.1 Workflow specification
1.2 Workflow execution in distributed systems
1.3 Monitoring of long-running workflows
1.4 Adapting components to the framework

2 Metadata, data description and logical organization

2.1 Data Format and Model Description
2.2 Managing Metadata
2.3 Using Data Descriptions and Relationships

3 Efficient access and queries, data integration

3.1 Large-scale feature-based Indexing
3.2 Query processing over files
3.3 Data integration

4 Distributed data management, data movement, networks

4.1 Data Placement
4.2 Replica management and movement
4.3 Dataflow between components
4.4 Multi-resolution data movement
4.5 Networking with embedded storage/computation
4.6 Security – authorization for data access and resource usage, integrity

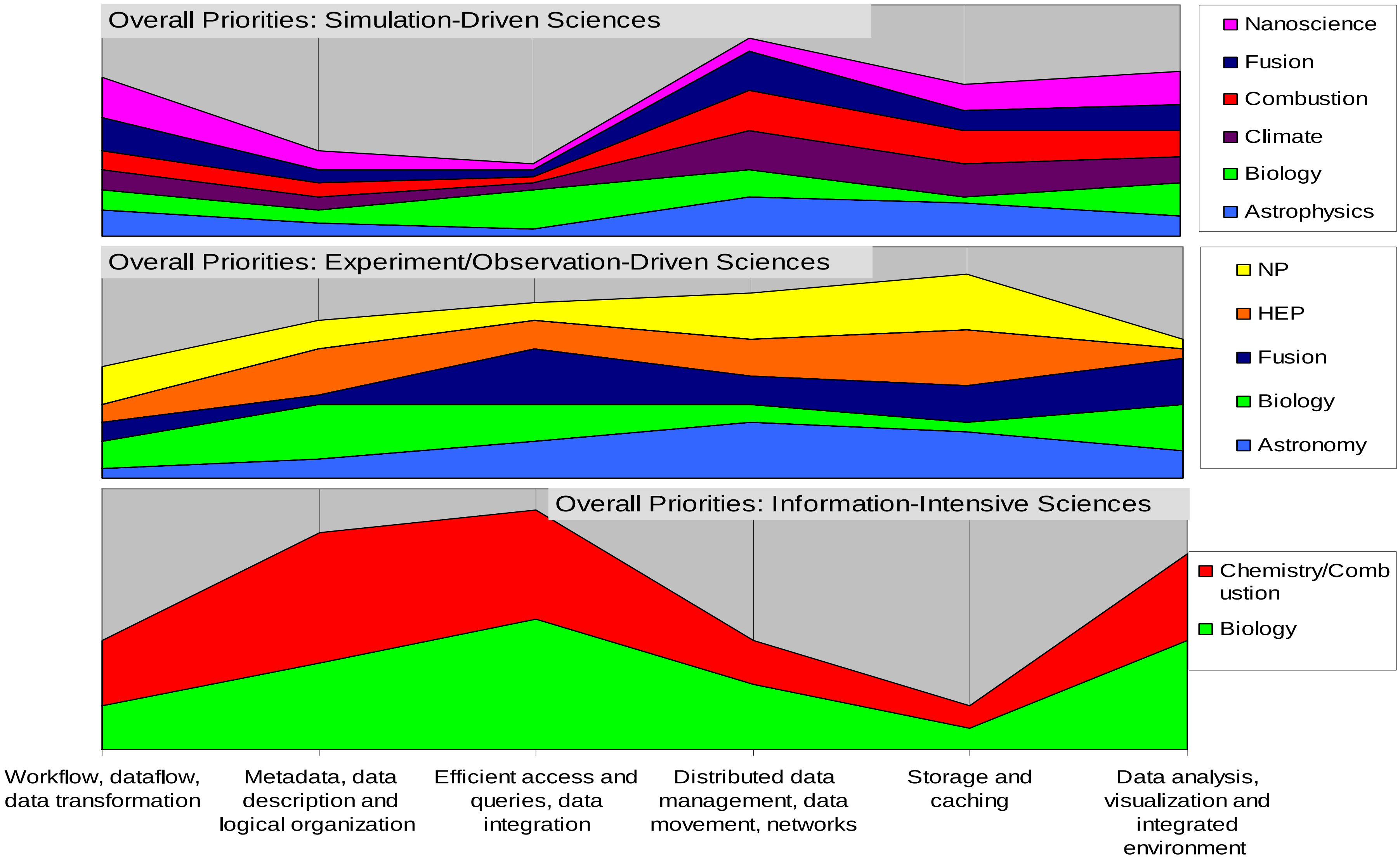
5 Storage and caching

5.1 Storage Technology
5.2 Parallel I/O: High-performance data access for computational science
5.3 I/O Access Patterns: Tuning systems for large working sets
5.4 Dynamic data storage and caching

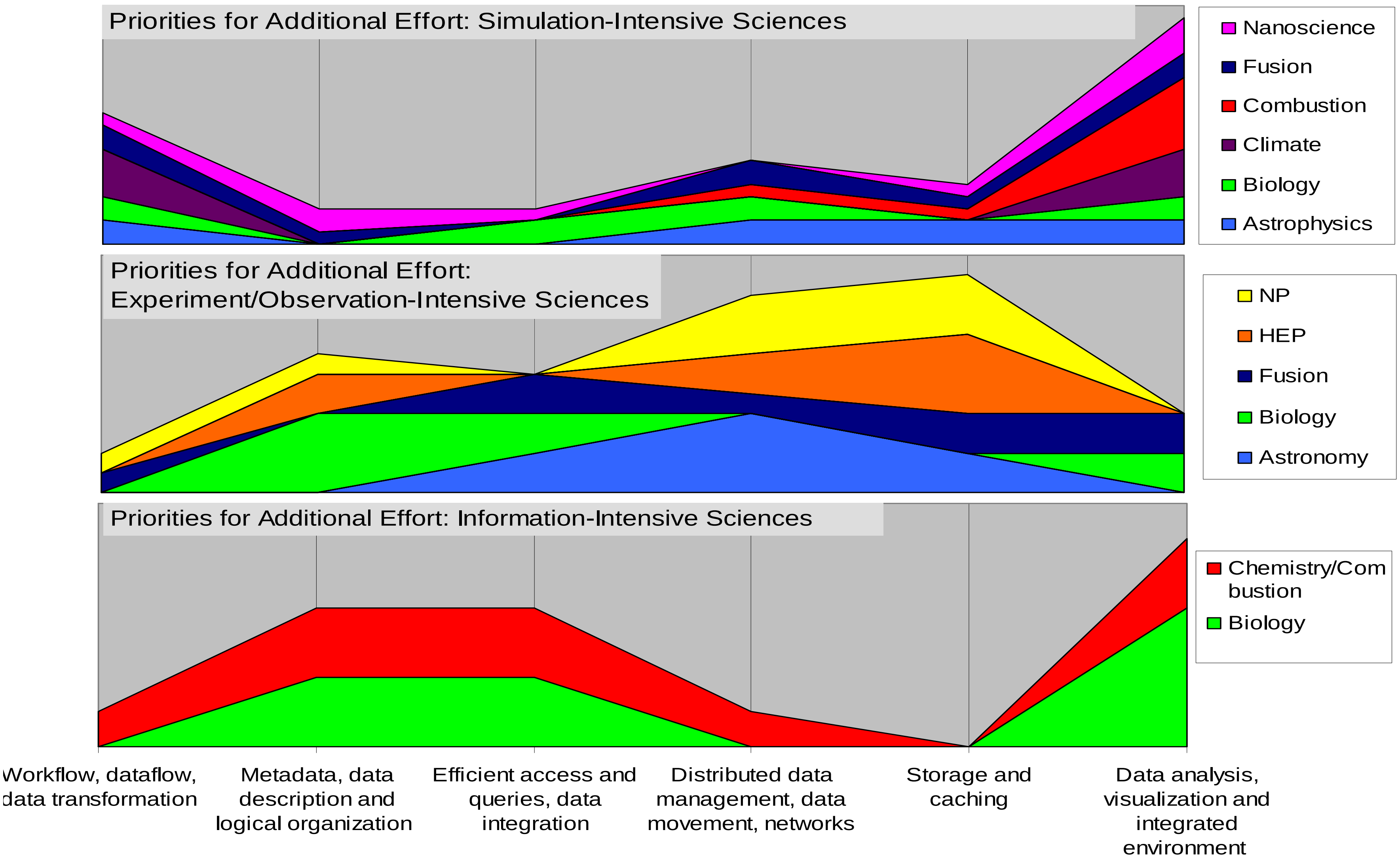
6 Data Analysis, Visualization, and Integrated Environments

6.1 Data Analysis
6.2 Visualization
6.3 Integrated Environments

Priorities


Overall Priorities: Simulation-Driven Sciences
Nanoscience
Fusion
Combustion
Climate
Biology
Astrophysics
Overall Priorities: Experiment/Observation-Driven Sciences
NP
HEP
Fusion
Biology
Astronomy
Overall Priorities: Information-Intensive Sciences
Chemistry/Combustion
Biology
Workflow, dataflow, data transformation
Metadata, data description and logical organization
Efficient access and queries, data integration
Distributed data management, data movement, networks
Storage and caching
Data analysis, visualization and integrated environment

Priorities for Additional Resources



Observations (1)

- Application sciences may be grouped as:
  - Simulation driven
  - Experiment/Observation driven
  - Information intensive
- Sciences within a group have very similar needs and priorities;
- Inter-group differences are striking;
- Every computer-science area is high-priority for at least one group.

Observations (2)

- Application sciences fund most of the data-management R&D;
  - Big science is surviving, but faces increasing difficulties
  - Small science is drowning.
- The CS-funded effort is well-regarded, but:
  - It is far too small
  - Hardening+packaging, support+maintenance are almost unrewarded and therefore dangerously inadequate.
- CS-application collaboration is productive.

Recommendation

An Office of Science Data-Management Program

- About $50M/year initially:
  - Half is new money
  - Half is from existing application and CS funding.
- Program Director
  - Align the program with evolving Office of Science needs
  - Ensure continuing buy-in by application science programs
- Guidelines for Proposals
  - Application-CS collaboration encouraged
  - Application-Application collaboration encouraged
  - Application-science contribution to funding is a major validation of the priority of the work.
- SciDAC is a good model.